Available online at www.sciencedirect.com

SCIENCE DIRECT

ELSEVIER

Vision Research 46 (2006) 2757–2772

Vision Research

www.elsevier.com/locate/visres

Flash-lag chimeras: The role of perceived alignment in the composite face effect

Beena Khurana [a,*], R. McKell Carter [b], Katsumi Watanabe [c], Romi Nijhawan [a]

[a] *University of Sussex, UK*

[b] *California Institute of Technology, USA*

[c] *National Institute of Advanced Industrial Science and Technology, Japan*

Received 4 November 2005; received in revised form 16 January 2006

Abstract

Spatial alignment of different face halves results in a configuration that mars the recognition of the identity of either face half (Young, Hellawell, & Hay, 1987). What would happen to the recognition performance for face halves that were aligned on the retina but were perceived as misaligned, or were misaligned on the retina but were perceived as aligned? We used the 'flash-lag' effect (Nijhawan, 1994) to address these questions. We created chimeras consisting of a stationary top half-face initially aligned with a moving bottom half-face. Flash-lag chimeras were better recognized than their stationary counterparts. However when flashed face halves were presented physically ahead of moving halves thereby nulling the flash-lag effect, recognition was impaired. This counters the notion that relative movement between the two face halves per se is sufficient to explain better recognition of flash-lag chimeras. Thus, the perceived spatial alignment of face halves (despite retinal misalignment) impairs recognition, while perceived misalignment (despite retinal alignment) does not.

Keywords: Face recognition; Facial form; Flash-lag effect; Configural encoding; Movement filter

1. Introduction

Human beings are exceptionally capable of recognizing individual faces (Bruce & Humphreys, 1994). The challenge of reliably individuating faces is made apparent by the fact that all faces share a basic configuration. Every individual face consists of facial features such as eyes, nose, and a mouth that have the same first-order relations such as two eyes above a nose and mouth (Maurer, Le Grand, & Mondloch, 2002). Although these features are most adequate in rendering the percept of 'a' face, they rarely render a percept of 'that' face (Liu, Harris, & Kanwisher, 2002). It has been suggested that the efficacy of face coding for the purposes of recognition must exploit second-order relational properties e.g., the spacing among the various features, over and above the features per se (Diamond & Carey, 1986; Liu et al., 2002; Maurer et al., 2002; Rhodes, 1988; Rhodes, Brake, & Atkinson, 1993; Sergent, 1984; Tanaka & Farah, 1993). These relational differences though small and undoubtedly requiring greater computational resource are thought to be sufficiently differentiable for accurate recognition.

Empirical evidence that relational information is an integral part of face processing comes from many different sources. For example, one approach has relied on measuring recognition performance when relational information is interfered with or compromised (McKone, Martini, & Nakayama, 2001). Tanaka and Farah (1993), on the other hand report poor recognition of individual isolated facial features. Based on these findings they propose that faces are processed 'holistically' such that information about distinct facial features is indivisibly combined with information

* Corresponding author. Fax: +44 1273 678058.
E-mail address: beenak@sussex.ac.uk (B. Khurana).

0042-6989/$ - see front matter © 2006 Elsevier Ltd. All rights reserved.
doi:10.1016/j.visres.2006.02.001

about their configuration. In consonance with these findings it was found that altering facial configurations impaired memory for facial features (Tanaka & Sengco, 1997). Other alterations to facial configuration that impede recognition include presenting face strips in different depth planes that cannot be amodally completed by observers to form the coherent surface of an entire face (Nakayama, Shimojo, & Silverman, 1989), breaking up a face into face parts (Farah, Tanaka, & Drain, 1995), and horizontally misaligning face halves (Moscovitch, Winocur, & Behrmann, 1997).

For the present experiments we turn to the 'composite face effect' (Young et al., 1987) as a signature of configurations being key in face perception. The effect is based upon facial chimeras that consist of the top half-face of one individual and the bottom half-face of another (Young et al., 1987). The visual system appears to treat these facial chimeras as a facial gestalt such that when observers are specifically asked to report the identity of one half face, the exclusion of the other half requires effort and comes with a cost in terms of time and or accuracy. Young et al. (1987) demonstrated the perceived integrity of facial chimeras by the impediment observed in the identification of either component half-face (Fig. 1A). However, when the components were spatially misaligned (Fig. 1B) the composite face effect was greatly reduced as evidenced by faster recognition of the identity of each face half. These original findings are taken to indicate that since chimeras give rise to the immediate perception of a new identity rather than a summation of linearly decomposable constituent face halves, alignment of face halves results in the mandatory activation of configural processes.

The original and subsequent studies have effectively employed physical alignment and misalignment of face halves to investigate the nature of configural processes. However, it has been shown that configural processes are

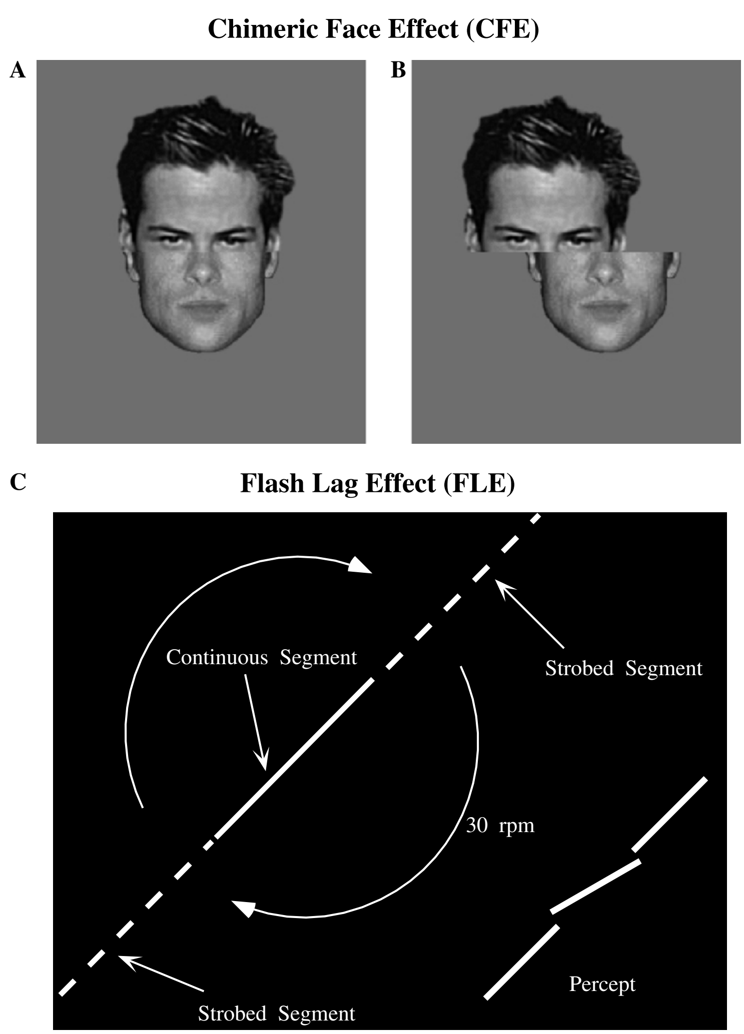

Fig. 1. (A) When the top and bottom half-face of two different individuals are aligned, the stimulus gives rise to a novel configuration that makes the recognition of either component half difficult. The phenomenon is referred to as the 'composite face effect' —CFE (Young et al., 1987). (B) When the top and bottom half-face of two different individuals are horizontally offset, the component halves are more readily recognized. (C) Two objects, one moving and one flashed briefly presented in spatial alignment give rise to the perception of flashed object as lagging the moving object. The phenomenon is referred to as the 'flash-lag effect'—FLE (Nijhawan, 1994).

sensitive to depth relations as well, thus making retinotopic brain areas unlikely sites for configural computations (Nakayama et al., 1989). Additionally, facial chimeras made of contrast reversed face halves also result in processing deficits (Hole, George, & Dunsmore, 1999). Thus, the configural processes implicated in the composite face effect are not contrast specific even though various behavioral studies have shown that contrast inversion mars recognition (Bruce & Langton, 1994; Galper, 1970; Hayes, Morrone, & Burr, 1986; Johnston, Hill, & Carman, 1992; Kemp, McManus, & Piggot, 1990; Phillips, 1972). Based on their findings, Hole et al. (1999) propose that facial chimeras may engage a more rudimentary form of configural processing, i.e., holistic processing that simply brings about the fast coupling of face features. Holistic processes can be distinguished from other configural processes that compute the relational aspects of features in order to identify a particular individual (Maurer et al., 2002). And it is configural processes that have been shown to falter when contrast is reversed. Given that holistic processes are fast acting, the time course of creating a facial gestalt is likely to be short (Lehky, 2000) in the chain of perceptual computations. In fact the most commonly held notion of Gestalt processes is one in which elements are grouped as a function of retinotopic features (Marr, 1982; Wertheimer, 1950); the suggestion is that processes before constancy is achieved are responsible for the observed grouping.

The configural processes responsible for the findings reported above are not fully specified. Here, we ask what would happen to the recognition performance for face halves that were aligned on the retina but were perceived as misaligned, or were misaligned on the retina but were perceived as aligned? In order to do so, we dissociate the physical configuration of the stimuli from the perceived configuration by devising facial chimeras based on the flash-lag effect (Nijhawan, 1994, 2002). When a moving and a flashed stimulus are presented in spatial alignment a compelling spatial dissociation between the physically given stimulus and the perceived stimulus occurs; namely the flashed stimulus is seen to spatially lag the moving stimulus (Fig. 1C). A variant of this procedure, one in which the trial is initiated by the flashed stimulus presented simultaneously with the moving stimulus renders a flash-lag effect that is comparable in magnitude to the standard complete cycle display, when the moving stimulus is presented both before and after the flash (Khurana & Nijhawan, 1995; Khurana, Watanabe, & Nijhawan, 2000a; Nijhawan, 1992). In order to address our current questions, we adapted this 'flash-initiated' variant to present two face halves; one in motion and the other flashed (Fig. 2). In this flash-initiated display, while the flashed and the moving items (face halves) are onset simultaneously they are displayed for unequal durations.

We had a second motivation to conduct the present experiments. In previous research it has been shown that the spatial offset observed in the flash-lag effect can have consequences and produce effects that are a by-product of spatial offset. For example, perceived spatial separation between two colored items, despite retinal co-location, can interfere with the 'mixing' of the two colors (Nijhawan, 1997). A similar question has not been experimentally

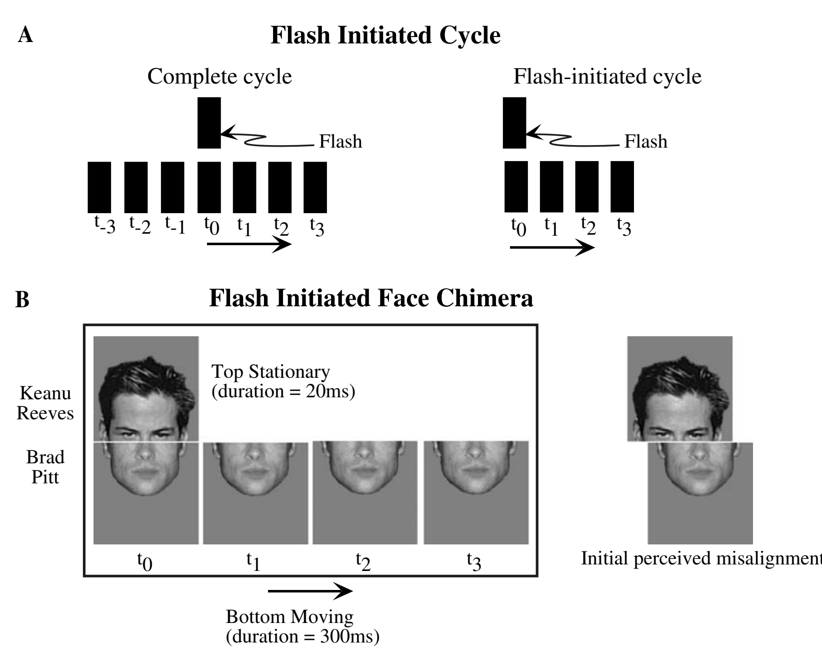

Fig. 2. (A) The flash-lag effect has been investigated using different cycles (Nijhawan, 1992). The complete cycle has the moving object visible both prior to and after the flashed object is presented. However, in the flash-initiated cycle the moving object is visible simultaneous with the flash and afterwards. (B) Figure shows the flash-initiated cycle for the face chimera stimuli. The top-half face is presented for only 20 ms simultaneous with the onset of the moving bottom half-face in frame t_0. In frame t_1 the top-half face is no longer visible which the bottom-half face has shifted to the right and continues to do for 280 ms. If the face stimuli are perceived as numerous other stimuli then the two faces would appear misaligned with the top-half face lagging the bottom-half moving face as shown to the right of the figure.

addressed for domains other than color. It would, for example, be of interest to ask the analogous question for shape. Our present experiments seek to answer one version of this question, namely whether spatial offsets can interfere with the computation of configurations in face processing.

In order to address these questions, we first had to establish a few facts concerning the composite face effect. First, we measured the identification performance on face halves presented briefly, and for unequal temporal durations. To this end, in Experiment 1 observers were presented a facial chimera, with the top half presented briefly (for one frame, see below) and the bottom half for a longer duration as either a static image or in a moving image. Thus in these variants the top half-face disappeared after the brief initial view while the bottom half remained visible (for an additional 14 frames). Observers were instructed to identify the top half of the face. We found that even under these limited and inequitable viewing conditions the perception of the top half face was affected by the presence of the bottom half. Next, we confirmed that when the top half of a facial chimera is presented in alignment with a moving bottom half, in a flash-initiated cycle, the top half is indeed seen as misaligned and lagging the bottom moving half. This was done in Experiment 2 where observers judged the relative positions of the flashed and moving face halves. Observers viewed face halves that were presented at various spatial offsets and reported whether the flashed top half appeared to lead or lag the moving bottom half in a two-alternative forced choice procedure (method of constant stimuli). From these data, we computed psychometric functions that provided a measure of how much in advance of the moving bottom half-face the flashed top half-face had to be presented in order for the two to be perceived in alignment. In Experiment 3, observers were asked to identify the flashed top half-face of a flash-lag face chimera while the bottom half was moving and misaligned relative to the top half. This misalignment constituted a forward shift (shift in the direction of motion) of the flashed half. The magnitude of the forward shift was determined by the results of Experiment 2; it was the point of subjective equality at which the flash was considered ‘ahead’ or ‘behind’ the moving item with equal probability. The key question was whether observers would be impaired at identifying the top half-face when the bottom half-face was perceived to be aligned, though it was physically misaligned relative to the top half. Additionally this experiment allowed us to measure the spatial offset at which reaction times peak, thereby allowing a determination of whether the lag/lead spatial judgments for that offset in Experiment 2 deviated significantly from 50%.

2. Experiment 1: Flash-lag chimeras

In order to gauge the recognition performance on a face halve in a flash-lag face chimera, we created a comparison control stimulus. In prior tests of recognition performance (Hole et al., 1999; Young et al., 1987) face chimeras have been presented to observers for durations such that they are visible on the screen until the observer responded. This would not be an appropriate baseline against which to compare the observer’s performance on flash-lag face chimeras in which the top half is briefly flashed while the bottom half either remains visible for a longer duration, or moves. Thus, Experiment 1 established whether the presentation condition, in which the two halves of facial chimeras are presented for brief and unequal durations, affects the observer’s identification performance on the top face half. The main goal of Experiment 1 was to find out if flash-lag based chimeras permitted more efficient access to face recognition processes by disabling the automatic activation of configural processes. We reasoned that if the face stimuli were perceived as other flash-lag stimuli, then observers would see one half-face spatially lagging the other. There were two possibilities. Were configural processes fast acting then perception of misalignment might not impact the configural processes triggered by the retinal alignment of the face halves. However, if input to configural processes were that of perceived alignment then flash-lag chimeras would not engage such processes (Fig. 3). We also presented flash-lag non-chimeras in which the top and bottom face halves belonged to the same individual (Fig. 4). Previously, we showed that while misalignment aids the recognition of chimera components, it has small costs for non-chimeras (Khurana, Watanabe, & Carter, 2000b).

2.1. Methods

2.1.1. Observers

Twelve observers (six male and six female, including authors BK, RMC, and KW) from the Caltech community volunteered to participate in the experiment. Observers had normal or corrected-to-normal vision. All observers except the authors were naive with respect to the hypothesis.

2.1.2. Apparatus and stimuli

Six famous male faces were presented for identification: The faces belonged to Brad Pitt, David Duchovny, Ricki Martin, Mel Gibson, Keanu Reeves, and Ben Affleck. Images were black and white frontal photographic stills downloaded from a website for celebrity pictures. Images were shown with a two-pixel gap between the upper and lower half faces because without a gap the moving bottom half-face appeared to cause some distortion at the boundary shared with the flashed top half-face. All the images showed famous males facing forward wearing a neutral expression. They were evenly lit and then adjusted for average brightness across the face using Photoshop. The images were also scaled so that all the faces were approximately the same height and width. However, none of these alterations impaired observers’ ability to recognize them. After the alterations the faces were 3.75° wide and 6.0° high. The faces were presented against a middle gray background in all experiments using a PowerMac with a 50 Hz monitor. Matlab along with the Psychophysics toolbox extensions

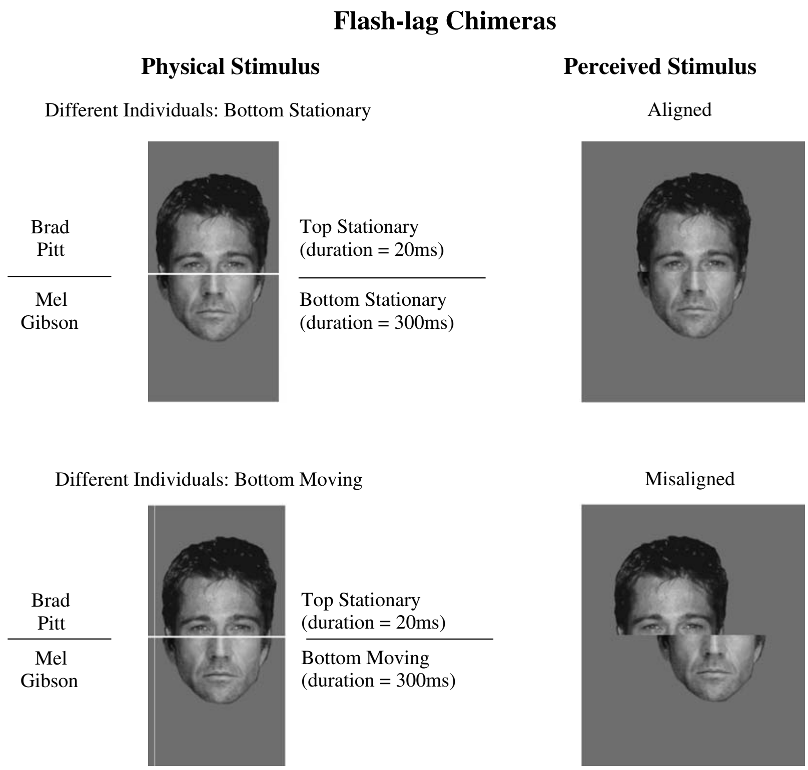

Fig. 3. Two kinds of flash-lag face stimuli were used in the three experiments. The first consisted of facial chimeras, i.e., face halves that belonged to different individuals (e.g., Brad Pitt and Mel Gibson) as shown in the figure. The top-half face was presented for 20 ms while the bottom-half face remained visible for 300 ms. Additionally, the bottom-half face was either stationary or moving from left to right at a speed of 12.5°/s. Both the physical and perceived stimuli are depicted.

(Brainard, 1997; Pelli, 1997) were used for stimulus presentation and data collection. Stimuli consisted of either aligned or misaligned face halves that were either taken from different individuals, i.e., facial chimeras (Fig. 3) or from the same individual, i.e., facial non-chimeras (Fig. 4).

On each trial chimeras or non-chimeras were displayed such that the bottom half-face was either stationary or moving. The resulting four trial types were randomly interleaved and each observer saw a different randomly generated sequence of trials. The top half face was presented for 20 ms and the bottom half face for 300 ms on each trial. During both the 'static' and the 'moving' trials, a complete face was presented on the screen for an initial period of 20 ms after which the top half disappeared and the bottom half either remained visible in the same location for an additional 280 ms or moved to the observer's right for 280 ms at 12.5°/s. The interval between the observer responding and the initiation of the subsequent trial varied from one to four seconds during which observers were instructed to maintain fixation. Six faces yielded 30 different facial chimeras. Four repetitions of each facial chimera were tested under both the moving and the stationary condition. Thus each observer was presented 48 trials of non-chimeras and 240 trials of facial chimeras.

2.1.3. Procedure

Before beginning the experiment, observers were trained on stationary complete faces of individuals. The purpose of this training was twofold. First, it was to establish whether observers could accurately identify the faces when presented briefly. Second, observers were provided an opportunity to establish a response mapping between the keys and the faces; each key corresponded to one of six famous faces mentioned above. Each face was presented for 20 ms and observers had to identify the individual faces with a key press. Observers were asked to place the index, middle, and ring finger of their right and left hands on six different keys on the keyboard. An accuracy of 95% or greater during the training session had be to achieved in order to start the experiment. Most observers required two training sessions of 100 trials each, in order to achieve the required level of competence. In the experiment, observers were asked to identify only the top half of the face stimuli presented on every trial and were informed that the top and bottom halves of a given face could belong to different individuals. In addition, they were asked to respond as quickly as possible while avoiding errors. Feedback was provided in the form of a high-pitched auditory beep for a correct response and a low-pitched beep for an incorrect response both during the practice session and main experiment. A complete session lasted approximately 15 min. Room illumination consisted of one overhead light.

During the experiment a white dot was centered on the screen between trials to allow observers to fixate prior to the presentation of a face. Reaction times and error rates

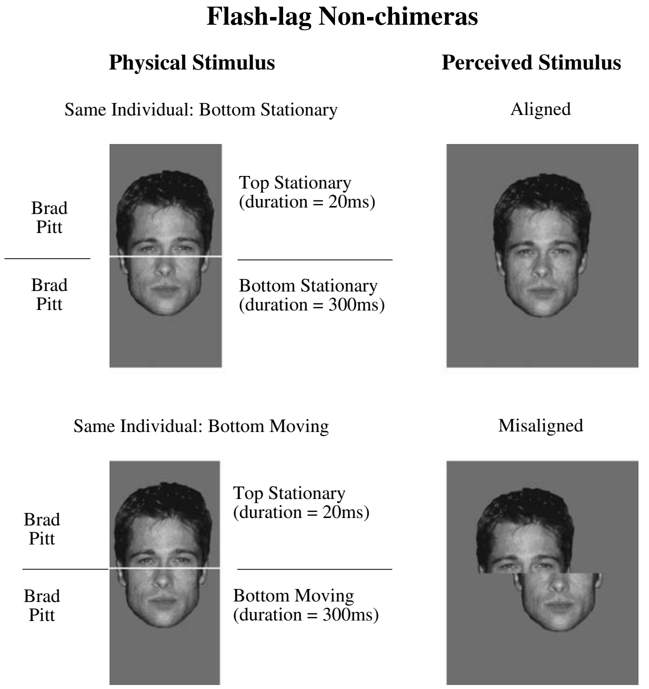

Fig. 4. Figure shows facial non-chimeras consisting of face halves that belonged to the same individual (e.g., Brad Pitt). The top-half was presented for 20 ms and the bottom-half (either stationary or moving) for 300 ms. Note that in the lower right-hand corner the perceived stimulus is similar to the physically misaligned stimulus used in the original investigation by (Young et al., 1987).

were measured. Reaction times were recorded using Matlab's built-in timer that is accurate to ±1 ms. A 'pass' key was also provided to allow an observer to not respond to a stimulus face that they did not feel capable of identifying accurately. The use of this key was very infrequent. However, for purposes of scoring, these responses were treated as incorrect answers. Observers responded by using the same keys that they were trained on. Observers were seated 57 cm from the monitor.

2.2. Results and discussion

A 2×2 (facial chimera versus non-chimera × bottom face halve stationary versus moving) repeated measures ANOVA on response times revealed a main effect of face type ($F(1,11) = 6.30$, $p < 0.03$). Thus, regardless of whether the bottom half was stationary or moving, observers were faster in the identification of facial non-chimeras as opposed to facial chimeras (mean reaction times for the non-chimeras and chimeras were 901.70 and 1037.10 ms, respectively; see Fig. 5A). Thus, though the components of the chimeras were presented for unequal durations, and the target top half-face was presented very briefly, the distractor bottom half-face did interfere with the identification of the top half-face. Whereas the main effect of movement was not significant ($F(1,11) = 2.72$, ns), there was a significant interaction between face type and movement ($F(1,11) = 4.57$, $p < 0.05$). Post hoc paired sample t tests showed that the identification of the top half of non-chimeras was not affected by whether the bottom half was stationary or moving (901.6 ms versus 901.8 ms; $t(1,11) = .01$, ns). But most critically for the present hypothesis the top half-face of facial chimeras are more readily identified when the bottom half face is moving (1006.5 versus 1067.6 ms; $t(1,11) = 2.61$, $p < 0.02$, Fig. 5C).

A similar 2×2 repeated measures ANOVA on accuracy indicated that observers were more accurate at identifying the top half of a facial non-chimera than a facial chimera (95.1% versus 82.2%, $F(1,11) = 4.77$, $p < 0.05$, Fig. 5B). No other effects or interactions were significant. The error data makes clear that the reaction time difference between flash-lag chimeras (82.0%) versus stationary chimeras (82.4%) was not due to a speed/accuracy trade off.

Observers found the task demanding as the half-faces they were required to identify were presented for a very brief duration. Most observers stated that they had to concentrate on the task to successfully respond. Nonetheless, under these restricted viewing conditions requiring a key press identification response—as opposed to unlimited viewing culminating in a vocal naming response in previous studies—we obtained differential performance under the static versus moving conditions. The top half of a flash-lag face chimera was more quickly identified than a face chimera in which the bottom half was stationary (see Fig. 5C). Thus, though the initial 20 ms of the static versus the moving trials were identical, the chimeras with a moving bottom half more easily permitted the identification of the flashed top-half. On the view that a flash-lag effect occurs in such displays, the accounting of the results is straightforward. We make the assumption that observers perceived a misalignment in the flash-lag based stimuli. This perceived misalignment was available to the holistic/configural process therefore facilitating recognition of the top half-face in the presence of a bottom half-face belonging to another individual.

Alternatively, one could reason that the static chimeras activated configural processes while the flash-lag chimeras did not, but not via the perception of misalignment. One possibility is that moving stimuli are treated distinctly in terms of processing from static stimuli. There have been previous suggestions of a 'neuroanatomical movement filter' that segregates moving stimuli from static stimuli (Cohen, 1999; McLeod, Driver, & Crisp, 1988). The site of the filter is thought to be the mid-temporal visual area (MT or V5) that responds well to moving forms but not static ones (McLeod, Heywood, Driver, & Zihl, 1989). The hypothesized movement filter is thought to act on global features or objects rather than local features. In the present instance the movement filter would be engaged by the moving bottom half of the flash-lag chimera. The findings of Experiment 1 can be considered analogous to previous findings of efficient search for a conjunction target defined by movement and shape (McLeod et al., 1988) as opposed to shape and color (Treisman & Gelade, 1980). Thus in the present experiments one can argue that attentional selec-

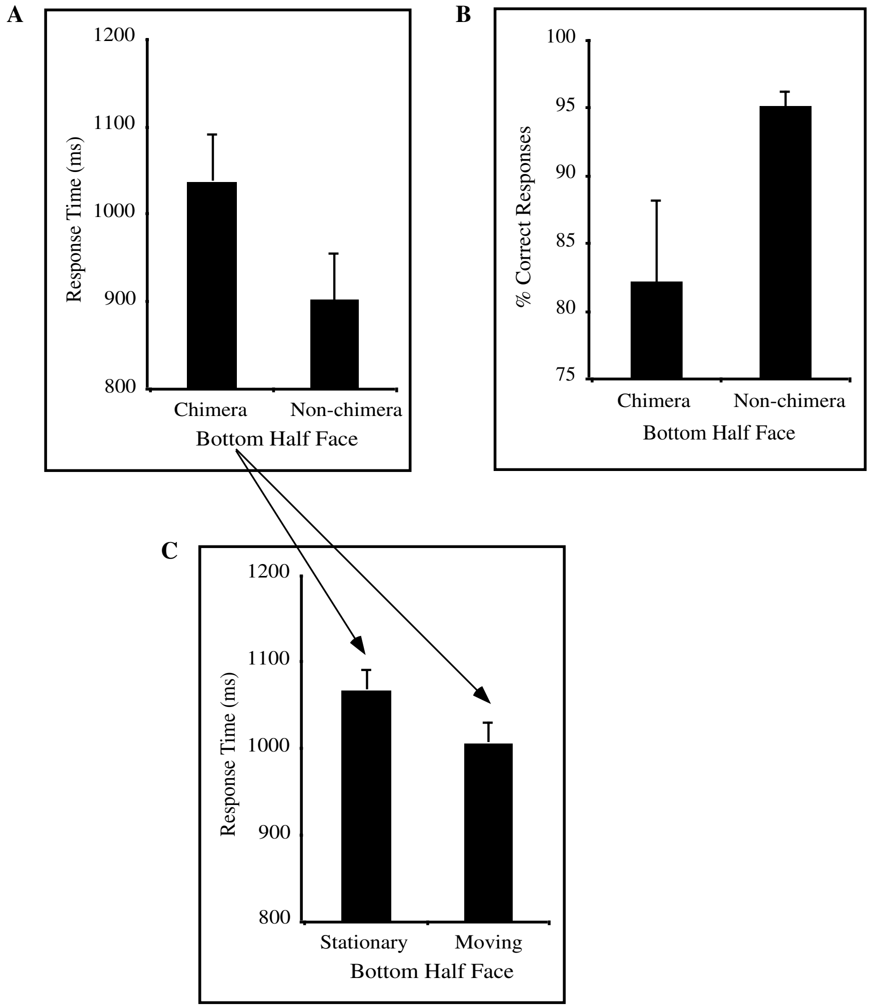

Fig. 5. (A) Chimeras are recognized more slowly than non-chimeras ($p < 0.05$). (B) Chimeras are recognized less accurately than non-chimeras ($p < 0.05$). (C) Chimeras are recognized more quickly when the bottom-half face is moving as opposed to being stationary ($p < 0.01$). Thus, when the stimulus conditions give rise to perceived misalignment even in the presence of physical alignment observers are faster to identify the components of a facial chimera.

tion through the movement filter permitted faster identification of the top half of the flash-lag chimeras. This account differs from the one offered by the flash-lag effect in that the flash-lag account endorses a perceived spatial offset as the mechanism through which recognition efficiency is achieved.

In this first experiment neither did we ask observers to report on the spatial aspects of their percepts nor measure them independently. This did not allow one to distinguish between accounts based on selective attention versus spatial offset via the flash-lag effect. While accounts of the flash-lag effect have included various retina-based mechanisms such as the persistence of the flashed item following its presentation (Nijhawan, 1992, 1994), and contrast gain control (Berry, Brivanlou, Jordan, & Meister, 1999), in this paper we focus on the perceived spatial effect per se and its consequences on face processing rather than on alternative accounts forwarded for the flash-lag effect. Experiment 2 was designed to address the issue of whether a flash-lag effect is indeed present for flash-lag chimeras.

3. Experiment 2: Flash-lag effect for facial chimeras and non-chimeras

In Experiment 1, we found that observers were faster to respond to flash-lag chimeras relative to static ones. We posited that these reaction time differences were due to a perceived spatial offset of the flashed half-face relative to the moving half-face. Such a percept is in line with a host of previous findings showing that flashed objects are perceived to lag physically aligned moving ones. But one might still wonder whether observers did indeed perceive the face halves to be misaligned. It could be conjectured that the moving bottom face provided less interference because after the initial 20 ms it did become physically misaligned. However, it is important to note that this misalignment would be between a representation/memory trace of the top half and a visual percept of the bottom half. Though, we solicited casual reports from the naive observers in Experiment 1 about the perception of misalignment, we did not get any quantitative measurements. In Experiment 2, we set out to explicitly measure the degree of perceived alignment when two face halves are presented, one moving and one flashed with varying degrees of initial offset. From the obtained localization data psychometric functions can be computed that render a point of subjective equality at which observers perceive alignment in the presence of physical misalignment in the stimulus. On the basis of the obtained data, we can explicitly test the motion versus spatial offset accounts (see Experiment 3).

Once again observers were presented with facial chimeras and non-chimeras. The top and bottom halves were presented for 20 ms and 300 ms, respectively. The bottom half was either aligned or misaligned to varying extents at onset. Observers were asked to judge the location of the top half face relative to the bottom face, by pressing 'ahead'/'behind' keys, in a two-alternative forced-choice procedure.

3.1. Methods

3.1.1. Observers

Four psychophysically trained observers (two male and two female) from the Caltech community participated in the experiment. Observers were required to have normal or corrected-to-normal vision. Observer CMG was naive as to the hypotheses being tested while authors BK, RMC, and KW were not.

3.1.2. Apparatus and stimuli

The apparatus was identical to that used in Experiment 1. Stimuli consisted of either aligned or misaligned facial chimeras and non-chimeras.

After the initial 20 ms the bottom half of the face was set in motion either to the left or right. During the initial 20 ms the bottom half-face was either aligned with the top half-face or misaligned by five different extents. The extent of misalignment increased in steps of 0.3°, with a maximum misalignment of 1.5°. Each offset was repeated 40 times. The facial chimeras and non-chimeras were tested in separate sessions. Each session consisted of 240 trials of varied offsets randomly interleaved.

3.1.3. Procedure

Observers were asked to place the index and middle finger of their right hand on the '1' and '2' keys of the number pad. Observers were instructed to press the '1' key to indicate that the top half-face was to the left of the bottom half-face and to press the '2' key if the top half-face was to the right of the bottom half-face. They were informed that responses were not timed. Observers were seated 57 cm from the monitor while they fixated a white dot that was centered on the face stimuli between trials.

3.2. Results and discussion

Observers found the task comfortable and the decision regarding offset easy. First, if observers did not perceive a flash-lag effect with these facial stimuli then one would expect responses to be centered on 50%, i.e., the flashed top half-face would be seen lagging or leading the bottom moving half-face equally often. However, observers showed a significant flash-lag effect in that they saw the top flashed half-face lagging the moving bottom half-face. Thus, when the face halves were physically aligned every observer deviated significantly from the 50% mark indicating a strong bias to perceive the faces as misaligned in the direction of the flash half-face lagging the bottom moving half-face. For the psychometric functions, an ANOVA analysis revealed a significant difference in the percent of 'flashed half ahead' responses at the various spatial offsets. Additionally, 'flashed half ahead' responses at zero spatial offset were close to zero (7.5% on average) while those at an offset of 1.5° were near 100 (92.8% on average) (Figs. 6 and 7). This indicates that the psychometric functions achieved an asymptote such that by an offset of 1.5° all observers perceived the flashed top half-face to be ahead of the bottom moving half-face. These data were fit using a logistic function $y = c/\{1+\exp[-a(x-b)]\}$. The point of subjective equality was calculated for each observer individually. It varied from 0.48° to 1.17° for different observers. This implies that observers in Experiment 1 indeed perceived the flash-lag facial stimuli as misaligned. The spatial mislocalization in which the flashed half-face perceptually lags the moving half is in agreement with a large number of findings using various visual stimuli such as lines (Nijhawan, 1994; Purushothaman, Patel, Bedell, & Ogmen, 1988); dots (Baldo & Klein, 1995; Khurana & Nijhawan, 1995; Whitney, Murakami, & Cavanagh, 2000), colored bars and disks (Nijhawan, 1997; Sheth, Nijhawan, & Shimojo, 2000), ring and disks (Eagleman & Sejnowski, 2000; Khurana et al., 2000a), and geometric shapes (Watanabe, Nijhawan, Khurana, & Shimojo, 2001). In all instances, the flashed component has been perceived to lag the aligned moving component.

Intriguingly, the psychometric functions for the facial chimeras and non-chimeras differed consistently for all four observers, in that the flash-lag effect was smaller for the non-chimeras versus the chimeras for every single observer (Figs. 6 and 7, respectively). Thus at first pass it appears that the nature of input affects the perceived lag. Previously it has been shown that the flash-lag effect is asymmetric in that it is greater for the leading edge of a moving object relative to the trailing edge (Watanabe et al., 2001). Watanabe et al. (2001) propose that the interaction between the global configuration of moving objects and the representation of spatial position may provide a new and useful tool for the study of perceptual organization. Our present findings using face halves indicate that the processing of face halves that belong to the same familiar face can reduce the spatial lag. The naïve observer showed the smallest difference between the flash-lag effect for chimeras and non-chimeras. It may be that the additional exposure of the other observers to the faces used in the experiment might be responsible. At this point one can only speculate as to the cause of this reduction in the flash-lag effect. Perhaps this might reflect a grouping or categorization response. Alternatively, this could arise

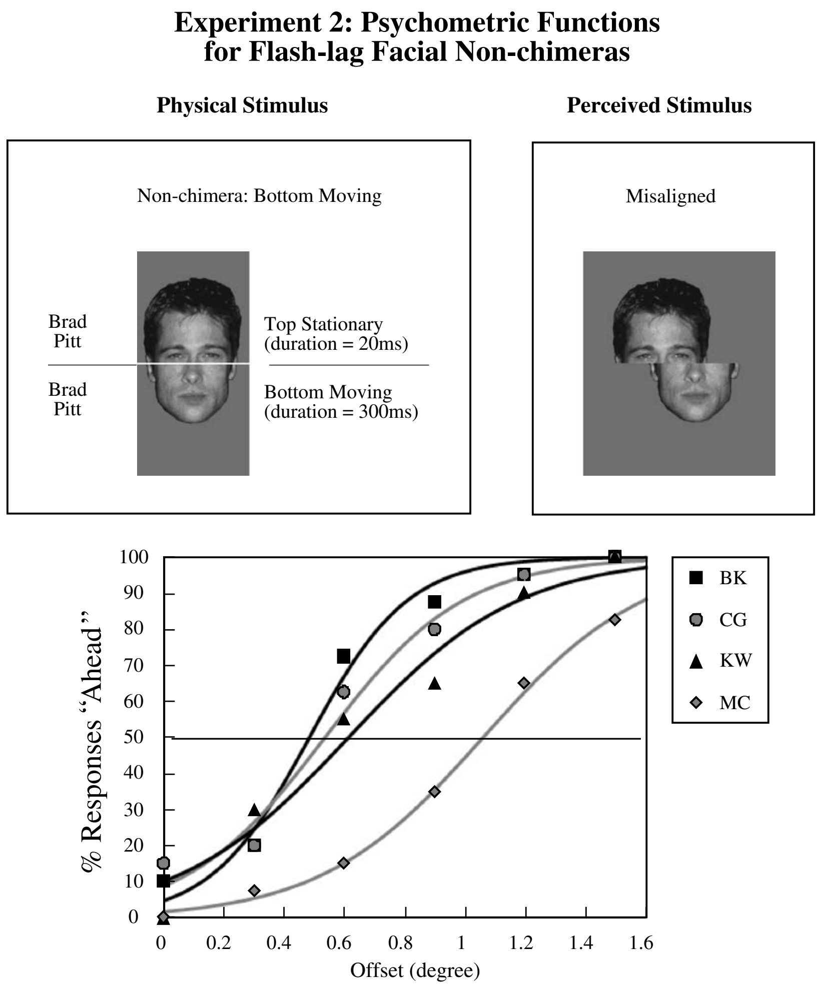

Fig. 6. Results of Experiment 2 in which observers were presented flash-lag non-chimeras and specifically asked to make a spatial judgment of whether the top-half face is 'ahead of' or 'lagging' the moving bottom-half face. Based on the responses curves were fitted using a logistic function $y = c/\{1 + \exp[-a(x - b)]\}$. The plots show that all four observers perceived the top-half face to be lagging the bottom-half moving face when the two were presented in spatial alignment (offset 0). The dashed line through 50% responses ahead is used to computed the point of subjective equality, i.e., the point at which the observer perceives the top-half face as being neither ahead of or behind the moving bottom-half face. This point varies as a function of the individual observer's psychometric function.

from priming in that the moving item is processed first and then it primes the processing of the temporally delayed flashed item. There may be some reduction in latency to the registration of the flashed item due to similarity to the moving item. Finally, salient or well-learned configurations might be capable of reducing perceived spatial offset when used in a flash-lag display[1]. We are currently following this finding up with a series of experiments testing how the nature of similarity between the moving and the flashed items affects the flash-lag effect.

4. Experiment 3: Speed of responses to flash-lag facial chimeras

In Experiment 3, we measured the observer's response time to identify the top half of either a facial chimera or non-chimera when the bottom half was moving and spatially offset relative to the top half-face to varying extents; thus we made direct measurements of response times to the very same spatial configurations used in Experiment 2. We were specifically interested in the comparison between reaction times collected for faces that were physically misaligned but were perceived to be aligned, against those for faces that were physically aligned but perceived to be misaligned. Data from Experiment 2 informs of when a given observer perceives two misaligned face halves as aligned. We asked: Would there be an increment in response times when the two components of a facial chimera, though physically misaligned are perceived to be aligned? Note that in Experiment 1, we ascribed the reduction in response times to a perceived spatial offset despite the physical alignment of the face halves, but there was an alternative possibility. The reduction in response times could be due to motion per se of the bottom face-half.

[1] The authors acknowledge J. López-Moliner for suggesting this bridge between ventral and dorsal processing.

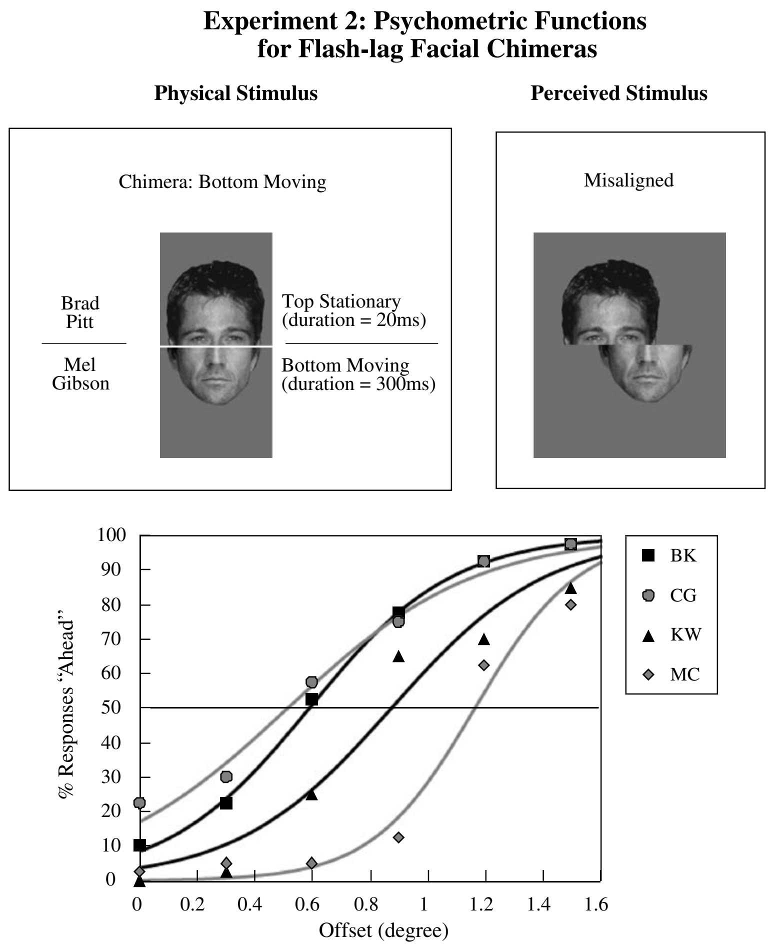

Fig. 7. Results of Experiment 2 in which observers were presented flash-lag chimeras and specifically asked to make a spatial judgment of whether the top-half face is 'ahead of' or 'lagging' the moving bottom-half face. Curves were fitted using a logistic function $y = c/\{1 + \exp[-a(x - b)]\}$. The plots show that all four observers perceived the top-half face to be lagging the bottom-half moving face when the two were presented in spatial alignment (offset 0 deg). The dashed line through 50% responses ahead intersects the curves at the point of subjective equality, i.e., the point at which the observer perceives the top-half face as being neither ahead of or behind the moving bottom-half face. These points vary as a function of the individual observer's psychometric function just as in the case of facial non-chimeras.

This experiment permitted a direct test of the movement filter account. According to motion filtering, the target top half-face is identified quicker in the presence of a moving bottom half-face because the observer can selectively filter out the moving component. If selective filtering is responsible for the decrement in reaction times to identifying the flashed top half-face then there should be no further modulation of response time as a function of different spatial offsets. Thus, the prediction from the movement filter account is that identification reaction times should be the same regardless of where the moving bottom half-face is presented relative to the top half-face.[2]

[2] Recall that throughout the range of offsets the bottom half-face after the initial 20 ms is always in motion. While spatial offset may be perceived, on the movement filter view this does not cause the reduction in reaction times to identifying the top half-face. According to filtering of attention via movement, as long as the bottom half-face moves, performance should be comparable.

4.1. Methods

4.1.1. Observers

The same four observers who participated in Experiment 2 took part in Experiment 3 in order to permit within observer comparisons between reaction times measured in this experiment and perceived alignments measured in Experiment 2.

4.1.2. Apparatus and stimuli

The apparatus and stimuli were identical to those used in Experiment 2.

4.1.3. Procedure

The experimental trials and sessions were identical to those in Experiment 2. The critical difference between Experiments 2 and 3 was in the responses made by the observers. Two separate sessions were run, one employing facial chimeras and the other non-chimeras. Observers were trained on the response keys used in Experiment 1 and were instructed to identify the top half of the face as quickly as possible without making errors.

4.2. Results and discussion

All four observers showed a peak in response times for the facial chimeras at a spatial offset different from zero (Fig. 8). A paired sample t test ($t(1,3) = 2.45$) showed these peak response times (Fig. 9A) to be significantly greater than response times at zero offset ($p < 0.05$). A similar analysis ($t(1,3) = 2.32$, $p < 0.05$) showed that accuracy was also compromised at offsets where peak response times were measured (Fig. 9B). Thus, not only were observers slower to recognize flashed face halves when they were physically misaligned, having been presented 'ahead' of the moving half, but they were also less accurate.

We then examined the data from Experiments 2 and 3. First, we took the spatial offsets at which the peak response times occurred in Experiment 3 and compared them with the point of subjective equality estimated from the perceived spatial offset task in Experiment 2. The offset that resulted in maximal interference from the bottom half-face was not significantly different from the estimated offset at which an individual observer perceived the flashed top half-face as neither lagging nor leading the bottom half-face ($t(1,3) = 0.21$, $p < 0.85$). However, one might object that we are comparing a discrete measure, i.e., a given offset at which reaction times peak in Experiment 3 with an estimated offset based on curve fitting the data in Experiment 2. Therefore, we also took the offset at which recognition performance was maximally affected in Experiment 3 and noted the exact percentage of lag reports for that very offset

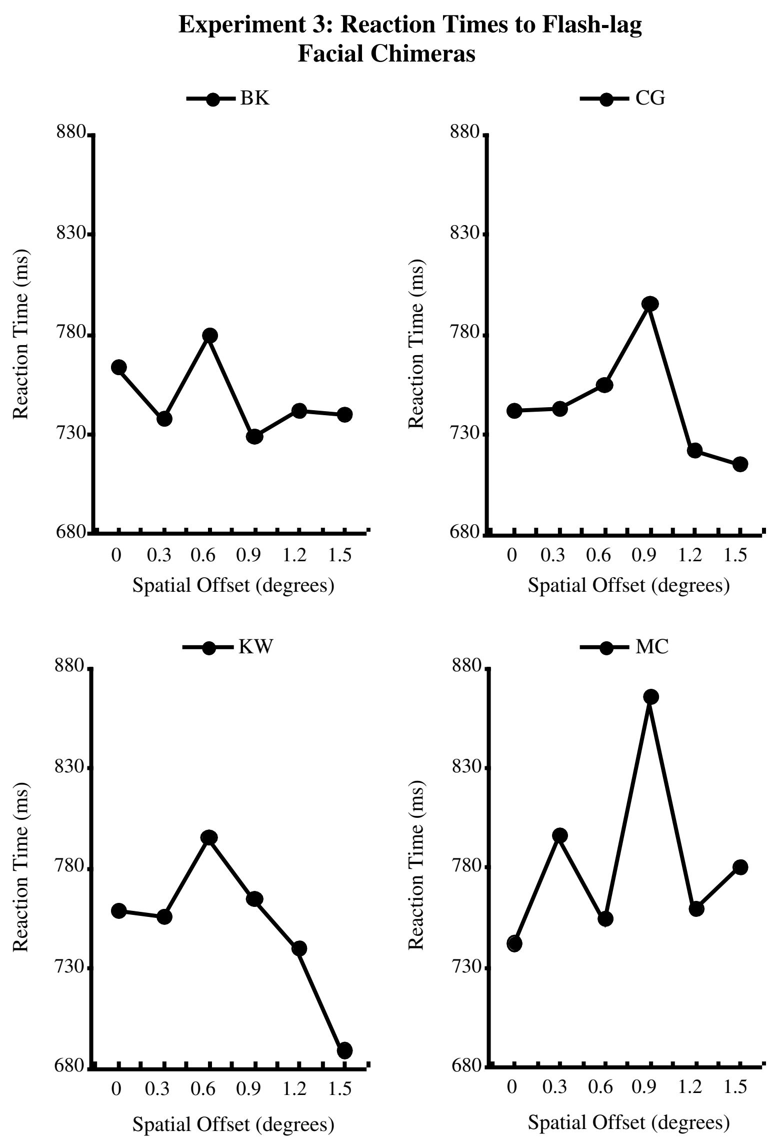

Fig. 8. Results of Experiment 3 in which observers were instructed to identify the top-half face as quickly and accurately as possible. Data is shown for the four observers that participated in Experiment 2 and only for flash-lag face chimeras. For every single observer the peak reaction time occurred at an offset greater than 0 deg and varied between 0.6 and 0.9 deg. Note that at these offsets the static top-half of the face chimera is being presented ahead of the moving bottom-half in the first frame.

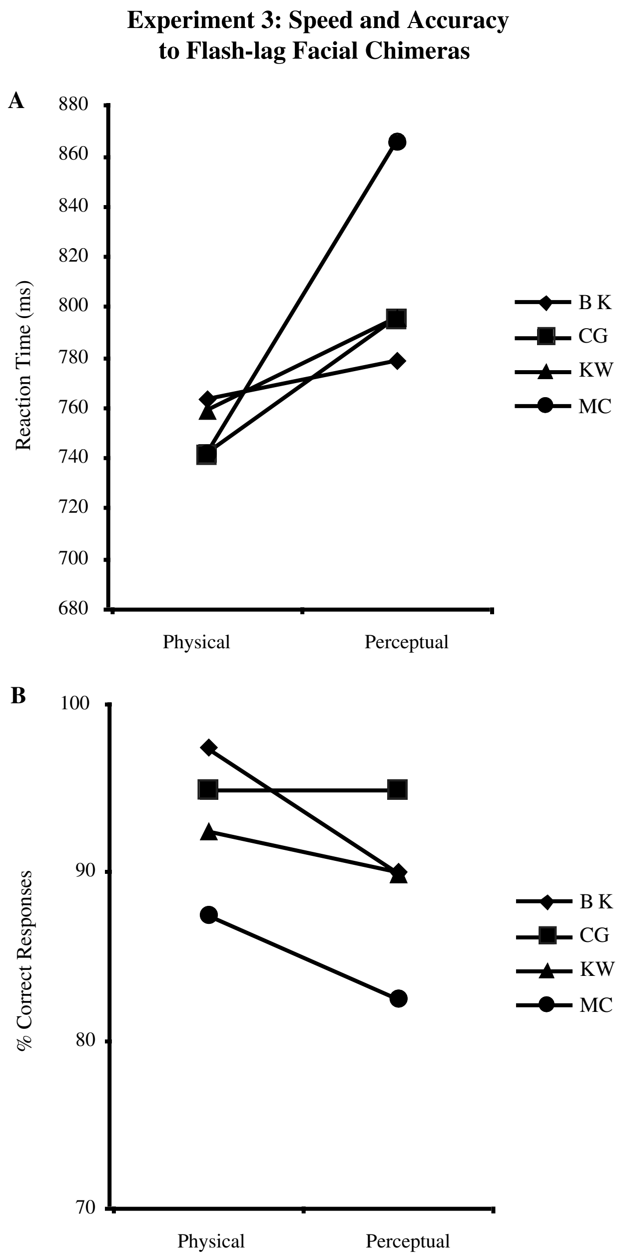

Fig. 9. (A) Reaction times to flash-lag chimeras peak at offsets greater than 0 deg. The plot shows the significant difference between reaction times at offset 0 (Physical Alignment) with peak reaction times (Perceptual Alignment). (B) The plot shows observers to be either more or equally accurate at offset 0 (Physical Alignment) than at the offset at which peak response times were measured (Perceptual Alignment). Thus, observers were not only slower but also generally less accurate when the two face halves were physically misaligned but perceived to be aligned.

in Experiment 2. Once again, we found no significant differences between the actual 'flashed half-face ahead' responses made at offsets where reaction times peaked with the null hypothesis value of 50% (point of subjective equality) ($t(1,3) = 0.62$, $p < 0.58$). The coincidence between these measures of recognition performance in Experiment 3 and perceived spatial offset in Experiment 2 lends further support to the hypothesis that the peak response times for flash-lag face chimeras are a function of perceptual alignment. Once again, as in Experiment 1, no such modulation of response times was present for the non-chimeras (Repeated measures ANOVA; Reaction times: $F(1,3) = 0.78$, $p < 0.40$; Errors: $F(1,3) = 1.00$, $p < 0.30$).

5. General discussion

When two face halves belonging to different individuals are aligned, the recognition of either component is impaired relative to when they are misaligned (Young et al., 1987). This composite face effect is thought to be a consequence of the automatic activation of configural face processes when the visual system is presented facial stimuli. In the present research, we devised a novel method to present chimeras such that perceptual alignment of face-halves could be decoupled from retinal alignment.

We report that the composite face effect can be observed with facial chimeras that consist of a briefly flashed top half-face and a longer duration bottom half-face, such that the two halves are initially aligned only for a brief duration. This shows a previously unknown robustness of the composite face effect. In Experiment 1 the bottom half-face was either stationary or moving. We found that observers were faster to determine the identity of the flashed half-face when the bottom half-face was moving as compared to when the bottom half-face was stationary. We hypothesized that this weakening of the composite face effect was due to the perception of misalignment caused by the flash-lag effect. In Experiment 2, we showed that observers indeed perceived the flashed top half-face to be lagging the moving bottom half-face in flash initiated displays. The results of Experiments 1 and 2 taken together indicate that the perceived spatial misalignment between the flashed and the moving face halves may cause the observed reduction in the composite face effect in Experiment 1. These findings suggest that configural face processes act on the output of processes that are responsible for either determining the movement status (motion versus static) or location of face halves. In Experiment 3, we found that observers were slowest to identify a face component that was retinally misaligned while being perceptually aligned thereby suggesting that differential movement of the two face halves did not contribute to a reduction in the interference offered by the bottom face half. Comparing the results of Experiments 2 and 3, we found that the degree of physical misalignment at which peak response times occur was not different from the point of subjective equality obtained from psychometric functions directly measuring perceived spatial offset.

Previous investigations of 'early' versus 'late' or pre-constancy versus post-constancy contributions to perceptual phenomena have argued for contributions from both levels (Palmer, Neff, & Beck, 1996; Rauschenberger & Yantis, 2001; Rock, Nijhawan, Palmer, & Tudor, 1992; Schulz & Sanocki, 2003). In the present context, we take pre-constancy phenomena to be based on relatively early computations representing sensory inputs and post-constancy phenomena to be based on later computations closer to perception (Treue, 2003). It has been suggested that post-constancy contributions are reflected only when observers have unlimited viewing time, while contributions from pre-constancy mechanisms are revealed when the viewing durations are limited (Rauschenberger & Yantis, 2001; Schulz & Sanocki, 2003). In the present experiments, although the presentation time of the flashed top half-face was limited to 20 ms, since we did not employ masking of the flashed stimulus the visual persistence would extend the visibility of the flashed component for about 100 ms (Hogben & Di Lollo, 1974). On the other hand, the duration of visibility of the bottom half-face in a given position would be restricted due to motion based de-blurring (Burr, 1980). Thus, the duration of representation in which the bottom half-face is spatially aligned with the flashed top half-face, will be shorter in the condition in which the bottom half is moving versus when it is stationary. This might account for the reduced composite face effect, despite the face halves being physically aligned in the two conditions.

It is worth noting that the differential persistence of the flashed and the moving elements may be offered not only as an account of the present effects, but of flash-lag itself (Krekelberg & Lappe, 2000; Nijhawan, 1992; Nijhawan, 1994). However, there are findings that oppose this interpretation. The flash-lag effect can be measured in the presences of masks that attenuate the persistence of the flashed item. Additionally, masking can be used to reduce persistence of a flash to see if the flash then behaves like the 'deblurred' moving item (Nijhawan, 1997). Both of these manipulations have been employed in the past. Whitney et al. (2000) presented flanking stimuli following the flash that acted as masks. They found that even when the visibility of the flashed item is limited by flanking stimuli, the flash-lag effect occurs undiminished. Second, Nijhawan (1997) showed that the 'color decomposition effect' does not occur when masking flanking bars restrict the duration of visibility of a flash; thus motion is necessary for the decomposition effect. This finding is consistent with dependence of other visual phenomena, such as acuity for apparent vernier offset, on visual motion (Burr, 1979). Such findings reinforce our suggestion that visual motion is necessary for the reduction of the composite face effect observed in Experiment 1. However, Experiments 2 and 3 suggest that visual motion per se is not sufficient and that the ensuing perceived spatial offset is necessary (see below). Thus it appears that perceived spatial alignment, whether based on retinal alignment or on motion (flash-lag effect), is necessary and sufficient for the composite face effect.

5.1. Movement based filtering

Previously it has been shown that movement permits the perceptual segregation of moving stimuli from static stimuli (Cohen, 1999; McLeod et al., 1988; McLeod et al., 1989). Rather than the flash-lag effect being responsible for disrupting configural processing, it could be the observer's ability to filter out the influence of moving objects that leads to the faster identification of the flashed top half-face. Experiment 3 permits a test of the validity of motion based filtering as an account of the reduction in the composite face effect. The movement based selective filtering account should predict equally fast recognition of the flashed top half-face whenever the bottom half-face is moving. The results of Experiment 3 showed this not to be the case. The composite face effect was rendered stronger or weaker as a function of perceived spatial offset despite the fact that the bottom half-face was moving in all trials. The flash-lag account predicts the observed increase in the composite face effect in the presence of a moving bottom half-face that is physically misaligned with the flashed top half-face. Therefore, though motion is necessary to cause the flash-lag effect, it does not appear to independently influence the magnitude of the composite face effect in the above experiments.

5.2. Percept–percept coupling

The approach we have adopted in the present experiments is related to the one adopted in the past by Rock and other investigators in which physical/sensory stimulation supports one percept while perceptual representations another. In this way the issue of whether a given phenomenon is based on 'early' versus 'late' processing has been addressed (Rock & Brosgole, 1964; Rock et al., 1992). In their classic study, Rock and Brosgole (1964) asked whether the Gestalt law of grouping by proximity was based on the anatomical closeness between the elements in the proximal stimulus or the closeness of the elements in perceived three-dimensional space. They manipulated physical versus perceived proximity and found that grouping substantially depended upon the perceived three-dimensional relation among the elements. The finding was characterized as that of one perception (three-dimensional space) influencing another perception (grouping of elements). Such outcomes have been thought of in terms of 'percept–percept coupling' (Epstein, 1982; Gogel & Koslow, 1972; Hochberg, 1974). In the present account, it is suggested that the perceived misalignment due to the flash-lag effect inhibits the action of configural processes thereby reducing the composite face effect.

Where does the interaction that gives rise to the above outlined percept–percept coupling take place? Since the misalignment caused by flash-lag is a directional effect, and as direction tuning in primates is mainly due to neurons in 'higher' visual areas (e.g., area MT/MST), our results imply an interaction between cortical motion processes and the composite face effect. In other words, we suggest that areas of the cortex that code for visual motion processing and the spatial localization of objects interact with those responsible for the recognition of faces. Previous experiments investigating the interaction between motion processing and face representation have reported that information in area MT/MST can influence the face processing area FFA via visual attention (O'Craven, Downing, & Kanwisher, 1999). O'Craven et al. (1999) posit that recurrent feedback from these extra-striate areas to earlier visual areas could enable such interactions. Related findings that argue for the late computation of configurations comes from MEG data in which the M100 is sensitive to face features whereas the M170 is more sensitive to configurations (Liu et al., 2002). Further support can be found in masking studies in which faces are best masked by upright faces regardless of differences in size, gender, and viewpoint (Loffler, Gordon, Wilkinson, Goren, & Wilson, 2005).

5.3. The 'what' 'where' crosstalk

In Experiment 2, we found the perceived spatial offset to be smaller when the component face halves belong to the same face as compared to when the component face halves belong to different faces for all observers. While this observation will require further investigation in order to establish its robustness it does support an account of the flash-lag effect in terms of processes located in higher levels of the visual pathway. If the flash-lag effect were based on 'early' processes then the similarity of face halves, which no doubt is computed by high-level identification mechanisms, could not impact the perceived spatial offset.

One account of the similarity based reduction in the flash-lag effect is that processing accorded to the moving segment makes contact with the underlying face representation and primes the matching of the flashed component (Khurana & Watanabe, 2001). This priming effectively leads to a reduction in the delay with which the flashed component is processed, and that in turn results in a smaller flash-lag effect. One may wonder if this reduction in the flash-lag effect will be present for other forms of similarity between the moving and the flashed halves, such as color, shape, texture etc. It may turn out that the reduced flash-lag effect for face halves belonging to the same individual has critically to do with the fact that half the face makes contact with the representation of the entire face, whereas such an argument is much more tenuous for more basic visual features such as colored segments. This may be because the visual processing related to the identification of faces is slower (Liu et al., 2002; Loffler et al., 2005) than that needed to process visual features, so the similarity between the halves can lead to a significant net reduction in the latency of processing one half. Thus, we suggest that the nature of a moving stimulus can have consequences on other processes such as the perception of a flashed item related in some manner to the moving stimulus. One consequence worth further exploration is that though unfamiliar faces give rise to the composite face effect (Hole, 1994), they should not result in significant differences in the flash-lag effect for chimeras made up of different unfamiliar face halves.

Such a modulation of flash-lag magnitude could be employed to define a continuum of 'relatedness' of objects or a continuum of object property constraints. Watanabe et al. (2001) previously showed that the global configuration of the moving stimulus affects the magnitude of the flash-lag effect. Based on those findings it was suggested that the flash-lag effect could be used as a tool to investigate perceived organization. More recently experiments using the flash-initiated cycle find that grouping occurs prior to the localization of moving and flashed stimuli (Watanabe, 2004). The present findings using facial chimeras along with others on grouping suggest that configural/organizational processes can impact the localization of objects.

More generally, such findings and their account have implications for how the visual system determines 'what' is 'where' (Ungerleider & Mishkin, 1982). It is now thought that representations in early cortical areas are dominated by sensory inputs gradually shifting to representations of perceptual interpretations at later cortical sites (Treue, 2003). Finding that early sensory computations of visual alignment can impact on later computations dealing with face processing is not surprising. However, our present findings suggest that representations at later cortical sites could impact on the coding of early sensory inputs. In Experiment 2, we found that the processing of what something is, i.e., a static top half-face belonging to the same individual as a moving bottom half face is localized closer than the top half-face of a different individual. Thus, it appears that later computations regarding facial identity presumably taking place in the inferotemporal cortex can impact early representations of spatial localization in the striate cortex. Additionally, though the dorsal and ventral pathways are specialized in terms of the visual functions they subserve, our present findings suggest that they are by no means independent.

5.4. Flash-lag chimeras and the flash-lag effect

In sum, we show that when the flash-lag effect occurs with face halves such that one is seen as misaligned from the other, the consequences of this misalignment are similar to those of retinal misalignment. While numerous experiments have been conducted on the flash-lag effect, this is only the second instance where the focus of the study is not mislocalization. Rather the focus is to show that the spatial offset observed in the flash-lag effect can have consequences for other perceptual properties. Earlier it was shown that retinally co-located red and green color patches, appearing separated due to the flash-lag effect, appear as red and green rather than yellow (Nijhawan, 1997). Here, we show that the perceived separation caused by the flash-lag effect causes a similar effect with facial configuration. Regardless of what view is taken on the causes of the flash-lag effect the goal of the present research was to show that the consequences of the perceived misalignment on visual processing can be as compelling as those of retinal misalignment.

Acknowledgments

The authors acknowledge an anonymous reviewer and J. López-Moliner for her insightful comments. Author B.K. thanks Tracy Anderson for her interest and support and Jim Parkinson for his input regarding psychometric functions.

References

Baldo, M. V. C., & Klein, S. A. (1995). Extrapolation or attention shift? *Nature (London), 378*, 565–566.

Berry, M. J., Brivanlou, I. H., Jordan, T. A., & Meister, M. (1999). Anticipation of moving stimuli by the retina. *Nature (London), 398*, 334–338.

Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vision, 10*, 433–436.

Bruce, V., & Humphreys, G. W. (1994). Recognizing objects and faces. *Visual Cognition, 1*, 141–180.

Bruce, V., & Langton, S. (1994). The use of pigmentation and shading information in recognizing the sex and identities of faces. *Perception, 23*, 803–822.

Burr, D. (1979). Acuity for apparent vernier offset. *Vision Research, 19*, 835–837.

Burr, D. C. (1980). Motion smear. *Nature, 284*, 164–165.

Cohen, D. (1999). Elements or objects? Testing the movement filter hypothesis. *Journal of Experimental Psychology: Human Perception and Performance, 25*, 348–360.

Diamond, R., & Carey, S. (1986). Why faces are and are not special: An effect of expertise. *Journal of Experimental Psychology: General, 115*, 107–117.

Eagleman, D. M., & Sejnowski, T. J. (2000). Motion integration and postdiction in visual awareness. *Science, 287*, 2036–2038.

Epstein, W. (1982). Percept–percept couplings. *Perception, 11*, 75–83.

Farah, M. J., Tanaka, J. W., & Drain, M. (1995). What causes the face inversion effect. *Journal of Experimental Psychology: Human Perception and Performance, 21*, 628–634.

Galper, R. E. (1970). Recognition of faces in photographic negative. *Psychonomic Science, 19*, 207–208.

Gogel, W. C., & Koslow, M. (1972). The adjacency principle and induced movement. *Perception & Psychophysics, 11*, 309–314.

Hayes, A., Morrone, M. C., & Burr, D. (1986). Recognition of positive and negative bandpass-filtered images. *Perception, 15*, 595–602.

Hochberg, J. (1974). Higher-order stimuli and inter-response coupling in the perception of the visual world. In R. B. McLeod & H. L. Pick Jr (Eds.), *Perception essays in honor of James J Gibson* (pp. 17–39). Ithaca, NY: Cornell, University Press.

Hogben, J. H., & Di Lollo, V. (1974). Perceptual integration and perceptual segregation of brief visual stimuli. *Vision Research, 14*, 1059–1069.

Hole, G. (1994). Configurational factors in the perception of unfamiliar faces. *Perception, 23*, 65–74.

Hole, G., George, P. A., & Dunsmore, V. (1999). Evidence for configurational processing of faces viewed as photographic negatives. *Perception, 28*, 341–359.

Johnston, A., Hill, H., & Carman, N. (1992). Recognizing faces: Effects of lighting direction, inversion, and brightness reversal. *Perception, 21*, 365–375.

Kemp, R., McManus, C., & Piggot, T. (1990). Sensitivity to the displacement of facial features in negative and inverted images. *Perception, 19*, 531–543.

Khurana, B., & Nijhawan, R. (1995). Extrapolation or attention shift? *Nature (London), 378*, 565–566.

Khurana, B., & Watanabe, K. (2001). Priming of faces from one half to the other. *Investigative Ophthalmology and Visual Science, 42*, 3927.

Khurana, B., Watanabe, K., & Nijhawan, R. (2000a). The role of attention in motion extrapolation: Are moving objects 'corrected' or flashed objects attentionally delayed? *Perception, 29*, 675–692.

Khurana, B., Watanabe, K., & Carter, R. M. (2000b). Configural face processes use high spatial frequencies. *Perception, 29*, 16.

Krekelberg, B., & Lappe, M. (2000). A model of the perceived relative positions of moving objects based upon a slow averaging process. *Vision Research, 40*, 201–215.

Lehky, S. R. (2000). Fine discrimination of faces can be performed rapidly. *Journal of Cognitive Neuroscience, 12*, 848–855.

Liu, J., Harris, A., & Kanwisher, N. (2002). Stages of processing in face perception: An MEG study. *Nature Neuroscience, 5*, 910–916.

Loffler, G., Gordon, G. E., Wilkinson, F., Goren, D., & Wilson, H. (2005). Configural masking of faces: Evidence for high-level interactions in face perception. *Vision Research, 45*, 2287–2297.

Marr, D. (1982). *Vision: A computational investigation into the human representation and processing of visual information*. San Francisco: W.H. Freeman.

Maurer, D., Le Grand, R., & Mondloch, C. (2002). The many faces of configural processing. *Trends in Cognitive Sciences, 6*, 255–260.

McKone, E., Martini, P., & Nakayama, K. (2001). Categorical perception of face identity in noise isolates configural processing. *Journal of Experimental Psychology: Human Perception and Performance, 27*, 573–599.

McLeod, P., Driver, J., & Crisp, J. (1988). Visual search for a conjunction of movement and form is parallel. *Nature, 332*, 154–155.

McLeod, P., Heywood, C., Driver, J., & Zihl, J. (1989). Selective deficit of visual search in moving displays after extrastriate damage. *Nature, 339*, 466–467.

Moscovitch, M., Winocur, G., & Behrmann, M. (1997). What is special about face recognition. Nineteen experiments on a person with visual object agnosia and dyslexia but normal face recognition. *Journal of Cognitive Neuroscience, 9*, 555–604.

Nakayama, K., Shimojo, S., & Silverman, G. H. (1989). Stereoscopic depth: Its relation to image segmentation, grouping and the recognition of occluded objects. *Perception, 18*, 55–68.

Nijhawan, R. (1992). Misalignment of contours through the interaction of the apparent and real motion systems. *Investigative Ophthalmology and Visual Science, 33*, 1415.

Nijhawan, R. (1994). Motion extrapolation in catching. *Nature (London), 370*, 256–257.

Nijhawan, R. (1997). Visual decomposition of color through motion extrapolation. *Nature (London), 386*, 66–69.

Nijhawan, R. (2002). Neural delays, visual motion and the flash-lag effect. *Trends in Cognitive Sciences, 6*, 387–393.

O'Craven, K., Downing, P., & Kanwisher, N. (1999). fMRI evidence for objects as the units of attentional selection. *Nature, 401*, 584–587.

Palmer, S. E., Neff, J., & Beck, D. (1996). Later influences on perceptual grouping: Amodal completion. *Psychonomic Bulletin & Review, 3*, 75–80.

Pelli, D. G. (1997). The videotoolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision, 10*, 437–442.

Phillips, R. J. (1972). Why are faces hard to recognize in the photographic negative? *Perception & Psychophysics, 12*, 425–426.

Purushothaman, G., Patel, S. S., Bedell, H. E., & Ogmen, H. (1988). Moving ahead through differential visual latency. *Nature, 396*, 424.

Rauschenberger, R., & Yantis, S. (2001). Masking unveils pre-amodal completion representation in visual search. *Nature, 410*, 369–372.

Rhodes, G. (1988). Looking at faces: First-order and second-order features as determinants of facial appearance. *Perception, 17*, 43–63.

Rhodes, G., Brake, S., & Atkinson, A. P. (1993). What's lost in inverted faces? *Cognition, 47*, 25–57.

Rock, I., & Brosgole, L. (1964). Grouping based on phenomenal proximity. *Journal of Experimental Psychology, 67*, 531–538.

Rock, I., Nijhawan, R., Palmer, S. E., & Tudor, L. (1992). Grouping based on phenomenal similarity of achromatic color. *Perception, 21*, 779–789.

Schulz, M. F., & Sanocki, T. (2003). Time course of perceptual grouping by color. *Psychological Science, 14*, 26–30.

Sergent, J. (1984). An investigation into component and configural processes underlying face perception. *British Journal of Psychology, 75*, 221–242.

Sheth, B., Nijhawan, R., & Shimojo, S. (2000). Changing objects lead briefly flashed ones. *Nature Neuroscience, 3*, 489–495.

Tanaka, J. W., & Farah, M. J. (1993). Parts and wholes in face recognition. *Quarterly Journal of Experimental Psychology, 46*, 225–245.

Tanaka, J. W., & Sengco, J. A. (1997). Features and their configuration in face recognition. *Memory & Cognition, 25*, 583–592.

Treisman, A., & Gelade, G. (1980). A feature integration theory of attention. *Cognitive Psychology, 12*, 97–136.

Treue, S. (2003). Climbing the cortical ladder from sensation to perception. *Trends in Cognitive Sciences, 7*, 469–471.

Ungerleider, L. G., & Mishkin, M. (1982). Two cortical visual systems. In D. J. Engel, M. A. Goodale, & R. J. Mansfield (Eds.), *Analysis of visual behavior* (pp. 549–585). Cambridge MA: MIT Press.

Watanabe, K. (2004). Visual grouping by motion precedes the relative localization between moving and flashed stimuli. *Journal of Experimental Psychology: Human Perception and Performance, 30*, 504–512.

Watanabe, K., Nijhawan, R., Khurana, B., & Shimojo, S. (2001). Global configuration of moving stimuli modulates the flash-lag effect. *Journal of Experimental Psychology: Human Perception and Performance, 27*, 879–894.

Wertheimer, M. (1950). Laws of organization in perceptual forms. In W. D. Ellis (Ed.), *A sourcebook of gestalt psychology* (pp. 71–81). New York: Humanities Press.

Whitney, D., Murakami, I., & Cavanagh, P. (2000). Illusory spatial offset of a flash relative to a moving stimulus is caused by differential latencies for moving and flashed stimuli. *Vision Research, 40*, 137–149.

Young, A. W., Hellawell, D., & Hay, D. C. (1987). Configural information in face perception. *Perception, 16*, 747–759.